УДК 004.93

С.В. Корнеев
Консорциум «BaltRobotics», Польша
ул. Три Липы, 3, г. Гданьск, 80-172

ОПЕРАЦИОННАЯ СИСТЕМА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: БАЗОВЫЕ ОПРЕДЕЛЕНИЯ

S.V. Kornieiev
BaltRobotics Sp.z.o.o., Poland
Trzy Lipy str.,3, Gdansk, 80-172

OPERATIONAL SYSTEM OF ARTIFICIAL INTELLIGENCE: THE BASIC DEFINITIONS

В статье кратко проанализировано состояние в области определений понятия «искусственный интеллект». Применительно к задаче разработки операционной системы искусственного интеллекта (ИИ) предложены: иерархическая структура взаимодействия, которая включает ИИ индивидуального типа и ИИ коллективного типа, при условии, что обе логических структуры базируются на физическом операционном носителе ИИ индивидуального типа; определение ИИ индивидуального типа; определение ИИ коллективного типа. Отличительной особенностью предложенных определений является то, что они в большей степени по сравнению с известными допускают моделирование с помощью известных методов и подходов.

Ключевые слова: Искусственный интеллект; определение интеллекта; индивидуальный интеллект; коллективный интеллект; иерархия.

The article briefly analyzes the current definitions of the concept "artificial intelligence". As applied to the development of the operating system of artificial intelligence (AIOS) it is proposed: the hierarchical structure of AIOS with the interaction of “AI-individual type” and “AI-collective type”, when both logical structures are based in the physical operating medium of “AI-individual type”; the definition of “AI-individual type”; the definition of “AI-collective type”. A distinctive feature of presented definitions is that they, if compared with known ones, allow the simulation using known methods and approaches.

Keywords: Artificial intelligence; intelligence definitions; individual intelligence; swarm intelligence; hierarchy.

Вступление

Несмотря на полувековую историю исследований в области Искусственного интеллекта и значительно более продолжительную историю психологии до настоящего времени не существует общепризнанного определения понятия «интеллект». Дело осложняется тем, что ученым, инженерам и программистам для организации работ по разработке искусственного интеллекта приемлемыми могут быть только те определения, которые содержат в себе понятия и механизмы, которые могут быть смоделированы в пределах разумных экономических и временных ограничений.

Постановка проблемы

Актуальная в настоящее время задача разработки операционной системы искусственного интеллекта, требует новых определений, которые могли бы быть положены в основу технических спецификаций. При этом задача моделирования данных спецификаций должна быть принципиально реализуема на основе имеющегося в настоящее время научно-технического задела.

Анализ последних исследований и публикаций

Легг и Хаттер в 2007 [1] году сформировали полезный сводный отчет об известных определениях интеллекта и на основании обобщения определили

интеллект как «способность некоторого «агента» достигать целей в широкой гамме внешней среды и условий». Сейчас это называют концепцией «оптимизационной мощности интеллекта».

Следствиями данного определения являет то, что Искусственный интеллект (как агент) обладает следующими свойствами:

- взаимодействует с внешней средой и может влиять на нее;
- способен получать успешные результаты в достижении некоторых целей;
- способен адаптироваться к различным средам и их изменениям.

В авторитетном Обзоре “AI100” [2] в качестве базового определения интеллекта приведена формулировка Нильса Нильссона (Nils J. Nilsson): «Искусственный интеллект» это деятельность посвященная тому, чтобы сделать машины умными, а интеллект, это качество, которое позволяет сущности функционировать надлежащим образом и с предвидением в своей среде".

Если продолжить анализ существующих определений, то мы обнаруживаем включение в определения следующих понятий с понятными семантическими связями, это: память, знания, опыт, понимание, рассуждение, воображение, заключение и ряд других. Фундаментальная работа [7] добавляет и основывается на понятиях «думать рационально» (“think rationally”) и «действовать рационально» (“act rationally”).

Недостатками существующих определений является их, скажем так, «не конструктивность», т.е. существующие в настоящее время определения не дают возможности технически структурировать работу по разработке Операционной системы искусственного интеллекта (далее AIOS).

Цель исследования

Целью данной статьи является публикация результатов автора по разработке базовых спецификаций AIOS в виде определений «*интеллекта индивидуального типа»*, *«интеллекта коллективного типа»* и иерархического характера их взаимодействия.

Изложение основного материала

Представленные в данной статье результаты были первоначально последовательно опубликованы в статьях [3,4].

По мнению автора, оправдано следующее определение.

«Индивидуальный интеллект» определяется наличием *осознанной границы* (в физическом смысле), в пределах которой определены и существуют следующие функции:

- функция гомеостаза (поддержание постоянства структуры и параметров в пределах границы);
- информационная функция (поддержание и обеспечение постоянного поиска и сбора информации, ее хранения, классификации, формирования абстрактных понятий, их иерархии и ассоциативных связей, прогнозирование развития окружающей среды, и собственного в ней состояния);
- функция воображения и планирования;
- функция изменения собственного состояния;
- функция воздействия на окружающую среду;
- целевая функция — монотонная векторная функция, отражающая степень «удовлетворенности» агента, «врожденной потребностью» которого является ее позитивная максимизация — «супер-задача».

Понятие «осознанная граница» предполагает, что *искусственный интеллект индивидуального типа*, имеет возможность получения от этой *границы* сигналов

идентифицированного *вторжения* в пределы данной границы, *касания* данной границы внешними объектами и средой, *повреждения* данной границы и *изменения* ее параметров под воздействием окружающей среды и вследствие внутренних процессов.

Данное определение *индивидуального интеллекта* было обсуждено в социальной сети linkedin в представительной группе «Applied Artificial Intelligence» («прикладной искусственный интеллект»). В дискуссии, в частности, приняли участие, известные ученые: Menno Mafait (Нидерланды: http://mafait.org/challenge/); Jean-Philippe de Lespinay (Франция: http://www.tree-logic.com/), и другие – всего около 20 исследователей из десятка ведущих стран мира. Протокол дискуссии составил несколько десятков листов с ним можно ознакомиться по приведенной ссылке [5].

Автор в своем определении использовал математический конструктивный поход, который заключается в том, что для доказательства существования объекта (математического) он «конструируется». Конечно, представленное «конструктивное» определение интеллекта математическим, по сути, еще не является и это еще не доказательство его валидности, - нельзя умозрительно доказать, что при предлагаемом структурном построении мы получим в результате «интеллект».

Некоторое формальное доказательство также пока невозможно, т.к. для него: 1) должны быть отмоделированы все приведенные функции; 2) этого все равно не хватит, т.к. *критерий интеллекта* – его «лакмусовая бумажка» - еще не создан.

Тем не менее, «конструктивное определение» AI автором предложено, оно позволяет двигаться в проектировании вперед и прошедшие обсуждения показали, что оно было встречено доброжелательно и оценено, как перспективное для развития [5].

Обсуждение подходов к моделированию функций, входящих в представленное определение, не относится к предмету данной статьи. Но автор может уверенно утверждать, что любой специалист в обсуждаемой области, как минимум, может представить свое видение моделей функций по представленному перечню. Из корреспондентов автора никто пока не предложил как-то изменить представленный список ни в сторону расширения, ни в сторону его сокращения. Но конечно апробация предложения должна продолжаться.

Предложенное определение основано на: 1) признании *иерархичности интеллекта*, как минимум, в звене {«индивидуальный интеллект» - «коллективный интеллект»} - данный факт сегодня не является общепризнанным в среде специалистов по AI, хотя биологи и зоологи уже доказали существование очень сложной адаптационной деятельности у насекомых, животных, рыб на уровне, который не присущ любой отдельно взятой особи из этих популяций; и 2) представленное выше определение относится только к «индивидуальному интеллекту», а предложение по определению интеллекта коллективного приведено ниже.

Для интеллекта индивидуального было высказано предположение, что семантически понятия «индивидуальный интеллект» и «инстинкт самосохранения» <u>тождественны</u> и появление двух понятий можно пояснить как предметно, – они возникли в разных разделах науки, - так и исторически, - они возникли в разное время. С точки зрения автора, принятие данной парадигмы сильно упрощает моделирование.

Обсуждая с автором феномен иерархичности интеллекта в группе linkedin “Applied Artificial Intelligence” («Прикладной искусственный интеллект») Dr. Dirk Hüske-Kraus (Германия, Штутгарт) высказал очень интересную гипотезу: «Если перед нами Интеллект – при этом не важно, как мы пришли к этому заключению, –

главное, чтобы оно было истинно, – то *интеллект* находится как внутри данного *интеллекта*, так и вне его при допустимости множественного отношения принадлежности».

Данное соображение представляется автору очень сильным и похожим на истину.

Теперь рассмотрим *интеллект коллективный*. В человеческий популяции «коллективный интеллект» соответствует понятию «интеллекта социального». Но это явление – социальный интеллект – представляется слишком сложным и многогранным для моделирования на данном этапе, поэтому автор скорее ориентируется на понятие «интеллект роя» (“swarm intelligence”), уже достаточно исследованное и описанное (см., например, фундаментальное исследование [6]).

Зачем при обсуждении базовых спецификаций AIOS нам нужно понятие коллективного интеллекта? Дело в том, что когда мы проектируем AI OS нам нужно принять решения: 1) по составу интерфейсов; 2) по принадлежности формируемых AI-роботом понятий к *индивидуальному* и *коллективному* уровням (с «наследуемыми понятиями» – все понятно, но с формируемыми в процессе обучения и на основании опыта – все далеко не так очевидно).

Если мы примем решение ограничиться только «индивидуальным уровнем», то:

1) как мы собираемся организовать эффективные действия группы AI-роботов? Наверное, никто не будет спорить, что обученная группа, например, воинское подразделение действует намного эффективнее «прямой суммы участников».

2) как мы собираемся вводить для AI-роботов понятие, если и не «Бог», скажем, «супервайзер»?

Необходимость введения «супервайзера» для AI-роботов не является очевидной, но на данном этапе такую целесообразность принципиально исключать нельзя. Интерфейс с «супервайзером» также принадлежит к уровню «коллективного интеллекта» при его создании. Таким образом, предусмотреть некоторый дополнительный процесс, который будет «отвечать» за коллективный интеллект в виртуальном AI-роботе достаточно просто.

В реальных AI-роботах это тоже возможно. При этом, т.к. у AI-роботов точно нет «6-го чувства», а есть только физически определенная система сенсоров, то «физический уровень» интерфейса будет сведен к этим сенсорам и их драйверам.

Понятно, что то, что мы пока интуитивно понимаем как *коллективный интеллект* – это «нечто», что не сводится к индивидуальным интересам AI-робота, как и в социальном смысле интересы группы часто противоречат интересам индивидуума.

Коллективный интеллект при определенных обстоятельствах может потребовать крайнего риска существования AI-робота, а может и его «самопожертвования». По какому механизму это может быть запрограммировано?

Прямая команда на исполнение – не подходит, т.к. она предполагает наличие «супервайзера», а мы рассматриваем более общий случай, когда его нет, – т.е. AI-роботы рассматриваются как *одно ранговая система* – одноранговая система намного более надежна. Можно предположить, что AI-роботы в группе будут, подобно социуму, приобретать некоторые дополнительные «веса/приоритеты». Причем именно «приобретать, а не «назначать».

Попробуем сформулировать определение для «AI коллективного типа».

Наследуя определение Легга-Хаттера (2007) для *интеллекта индивидуального* можно предложить следующее:

«*AI коллективного типа* проявляется в виде согласованного поведения группы экземпляров AI-роботов, когда члены группы, в общем случае, выполняют различные ролевые действия, для достижения целей группы рациональным образом в изменяющихся условиях внешней среды».

Теперь попробуем дать определение *AI коллективного типа* в конструктивном смысле.

«Коллективный интеллект может при определенных условиях возникать в группе особей, обладающих *Интеллектом индивидуального уровня*, и направлен на достижение следующих целей: (1) увеличение популяции всеми доступными методами: размножение, защита членов популяции от посягательств на жизнь и здоровье, увеличение продолжительности жизни и условий существования членов популяции, а также (2) организация совместных действий по увеличению *жизненных ресурсов популяции*».

Для возникновения явления *AI коллективного типа* необходимо одновременное выполнение следующих условий:

1) Наличие группы экземпляров AI-роботов.
2) Наличие наследуемого/обученного понятия *родства:* одно рангового или иерархического.
3) Наличие наследуемого/обученного понятия *группового интереса.*
4) Наличие наследуемого/обученного *процесса/плана действий* группы для достижения целей по условию 3.
5) Наличие наследуемого/обученного «стимула/наказания» (возможно нескольких) для принуждения членов группы к следованию приоритетам группы».

Опция: *AI-член группы* может формировать некоторый *приоритет*, который при «предъявлении/индивидуальном восприятии» позволяет отдавать команды членам группы, либо при обращении к группе ранг приоритета будет прямо влиять на количество членов группы, которые будут следовать выраженному указанию. Этот *приоритет* может быть или наследуемым (вводимым при производстве) или назначаемым, – т.е. вводимым супервизором данного AI-аппарата.

Приоритет – это процесс, который должен иметь врожденную/обученную векторную функцию формирования на основании взаимодействия с членами группы.

Данный «приоритет» влияет на распределение «благ/поощрений» - в авторском определении индивидуального интеллекта – это целевая функция, к максимизации которой «по определению» стремится AI-робот.

Процесс коллективного интеллекта в *интеллекте индивидуальном* может наследуемым образом содержать следующие «постулаты»:

– при угрозе членам группы AI-член обязан защищать группу/(ее членов) тем более, чем более совокупный приоритет членов группы, находящихся под угрозой, превышает его собственный;

– AI-член группы должен выполнять команды, поступающие от AI-члена с большим приоритетом тем более активно, чем более существует взаимный сравнительный дефицит приоритета;

– AI-член группы обязан добиваться исполнения своих команд членом группы с меньшим приоритетом.

«Постулаты» коллективного интеллекта блокируют индивидуальный интеллект тем более, чем больше дефицит приоритета члена группы по отношению к «соплеменникам».

Утверждение выше, что коллективный интеллект «может возникать при определенных условиях» связан некоторым образом с понятием «компактности

группы», и эту «компактность» довольно просто определить: она определяется дистанцией возможного взаимодействия членов группы с помощью доступной им генерации сигналов и возможностями их восприятия сенсорной системой.

Приоритет AI-робота в группе может быть «коллективно осознан», например, на основании:

– частоты (в смысле математической статистики) дублирования группой (ее частью) поведения некоторого «AI-индивидуума»;

– частоты выражения поддержки действиям «AI-индивидуума («вербальная поддержка», следование и др.);

– отсутствия противодействия AI-индивидууму со стороны членов группы;

– частоты выигрыша AI-индивидуумом некоторых «соревнований» в группе (возможно некоторое коллекционирование рейтингов проигравших особей).

Приоритет может быть и наследуемым (но это уже противоречие «одноранговости AI-популяции»).

Говоря о *Коллективном интеллекте* к вышеприведенному перечню его целей можно добавить еще два, которые не столь прямо относятся к функции сохранения вида, но все же к ней относятся, это:

- обучение/самообучение (в коллективе, в том числе «наставничество»);
- выявление лидеров.

На уровне физических интерфейсов AI-робота введение *коллективного интеллекта* никак не отражается: взаимодействие AI-робота с внешней средой, в том числе, с себе подобными, организовывается через имеющиеся средства восприятия информации.

Природа этих датчиков понятна, это сенсоры: звукового диапазона, радиодиапазона, видимого света, инфракрасного, ультрафиолетового, гамма-излучения. Кроме этого можно предположить наличие химических анализаторов, а может и синтезаторов, тактильные датчики давления и оценки текстуры поверхности, – это представляется как исчерпывающий перечень для случая AI-робота – реального физического устройства для работы в реальной физической среде.

Однако это не все варианты реализации! Возможны AI-роботы – программы в виртуальной среде. В этом случае его сенсоры – это программные интерфейсы. В этом случае ситуация во многом становится проще: виртуальная среда практически полностью стандартизована/специфицирована. Все объекты виртуальной среды «a-priori» математически определены за исключением вредоносного ПО, которое определено «a-posteriori».

В социуме и у животных коллективными целями являются: 1) размножение; 2) совместное добывание пищи/охота; 3) защита от хищников и природных угроз; 4) строительство жилищ; 5) миграция.

Коллективные цели в живой природе всегда связаны с сохранением вида. По аналогии с авторской формулировкой *Интеллекта индивидуального,* в основу которой положен *инстинкт самосохранения,* автор склонен в основу определения *Интеллекта коллективного* также положить *Инстинкт сохранения вида*. Отдельный вопрос – насколько далеко мы должны заходить в наделении AI-антропоморфными свойствами или бионическими аналогиями.

Выводы

В представленной статье предложены:

– иерархическое построение искусственного интеллекта в составе *интеллекта коллективного* и *интеллекта индивидуального* применительно к задаче разработки *операционной системы искусственного интеллекта*, при этом процессы коллективного интеллекта базируются на объекте интеллекта индивидуального; индивидуальный интеллект доминируется интеллектом коллективным по предложенному механизму;

– определение индивидуального интеллекта;

– определение коллективного интеллекта.

Индивидуальный интеллект предложено рассматривать тождественным понятием с *инстинктом самосохранения особи*.

В основу определения коллективного интеллекта положено понятие *инстинкта сохранения вида*.

Литература

1. Legg S., Hutter M. A collection of Definitions of Intelligence / Technical Report, IDSIA-07-07, 2007. – 12 p. [Электронный ресурс]. – Режим доступа: https://arxiv.org/pdf/0706.3639.pdf
2. One Hundred Year Study on Artificial Intelligence (“AI100”), Standford University, 2016. – 52 p. [Электронный ресурс]. – Режим доступа: https://ai100.stanford.edu/2016-report
3. Корнеев С. Искусственный интеллект: как и зачем? / «Сети и бизнес» (Украина), №2 (81), 2015. – с.72-79.
4. Корнеев С. Искусственный интеллект как иерархическая структура / «Сети и бизнес» (Украина), №3 (88), 2016. – с.80-86.
5. The Discussion in Linkedin Group - "Applied Artificial Intelligence" on the definition by S. Kornieiev: "Artificial Intelligence: Individual one, the definition of constructive type", DOI: 10.13140/RG.2.1.1078.5682, [Электронный ресурс]. – Режим доступа: https://www.researchgate.net/publication/301217412_The_Discussion_in_Linkedin_Group_-_Applied_Artificial_Intelligence_on_the_definition_by_Sergii_Kornieiev_Artificial_Intelligence_Individual_one_the_definition_of_constructive_type
6. Bonabeau E., Dorigo M., Theraulas G. Swarm Intelligence: from Natural to Artificial Systems / New York, Oxford, Oxford University Press, 1999. – pp.307.
7. Russel S., Norvig P. Artificial Intelligence: A Modern Approach, 3-d Edition / Prentice Hall, New Jersey, 2009. – pp.1152. [Электронный ресурс]. – Режим доступа: http://aima.cs.berkeley.edu/

Literatura

1. Legg S., Hutter M. A collection of Definitions of Intelligence / Technical Report, IDSIA-07-07, 2007. – 12 p. [Elektronnyiy resurs]. – Rezhim dostupa: https://arxiv.org/pdf/0706.3639.pdf
2. One Hundred Year Study on Artificial Intelligence (“AI100”), Standford University, 2016. – 52 p. [Elektronnyiy resurs]. – Rezhim dostupa: https://ai100.stanford.edu/2016-report
3. Korneev S. Iskusstvennyiy intellekt: kak i zachem? / «Seti i biznes» (Ukraina), #2 (81), 2015. – s.72-79.
4. Korneev S. Iskusstvennyiy intellekt kak ierarhicheskaya struktura / «Seti i biznes» (Ukraina), #3 (88), 2016. – s.80-86.
5. The Discussion in Linkedin Group - "Applied Artificial Intelligence" on the definition by Sergii Kornieiev: "Artificial Intelligence: Individual one, the definition of constructive type", DOI: 10.13140/RG.2.1.1078.5682, [Elektronnyiy resurs]. – Rezhim dostupa: https://www.researchgate.net/publication/301217412_The_Discussion_in_Linkedin_Group_-_Applied_Artificial_Intelligence_on_the_definition_by_Sergii_Kornieiev_Artificial_Intelligence_Individual_one_the_definition_of_constructive_type
6. Bonabeau E., Dorigo M., Theraulas G. Swarm Intelligence: from Natural to Artificial Systems / New York, Oxford, Oxford University Press, 1999. – pp.307.
7. Russel S., Norvig P. Artificial Intelligence: A Modern Approach, 3-d Edition / Prentice Hall, New Jersey, 2009. – pp.1152. [Elektronnyiy resurs]. – Rezhim dostupa: http://aima.cs.berkeley.edu/

RESUME

S. Kornieiev

Operational System of Artificial Intelligence: basic definitions

The author in the article analyzed the current definitions of artificial intelligence and pointed out that they lack of base data for real design of "AI-operational system" (AI-OS).

It had been proposed:

–AI hierarchical structure composed of the *collective intelligence* and *individual intelligence*, wherein the *collective intelligence process* is based on the object of *individual intelligence*; and *individual intelligence* is dominated by a *collective intelligence* on the proposed mechanism;

–the definition of individual intelligence based on the concept of the *instinct of individual self-preservation*;

–the definition of collective intelligence based on the concept of *preservation of species instinct*.

The Definition of Individual Intelligence.

The Individual Intelligence has been defined by self-controlled "border" (in physical sense), and in the range of it the following functions are defined and exist:

–function of homeostasis: maintaining the constancy of structure and parameters within the boundary;

–information function: maintaining and ensuring the permanent collection of information; its storage; classification; formation of abstract concepts, their hierarchy and associative links; forecasting of the development of the environment and its own state;

–the function of imagination and planning;

–function to change its own state;

–the function of the environmental impact;

–an objective function - monotone vector function, reflecting the degree of "satisfaction" of agent - "inherent need" of which is to maximize its positive value – the "super-task".

The concept of "self-controlled border" implies that the *artificial intelligence of the individual type* has the ability to receive from this border identified invasion signals, signal of touch by external objects and the environment, signals of damage of the border and change its parameters under the influence of the environment and due to internal processes.

The Definition of Collective Intelligence.

«The Collective Intelligence can occur in the group of AI-robots with AI of Individual Type, and try to reach the following objectives: (1) to enlarge the population by all controlled means: reproduction, protection of members of the population from attacks on life and health, increase life expectancy and living conditions of members of the population, and (2) to organize mutual activity of members to enlarge the living recourses for population».

For the occurrence of the phenomenon of the AI Collective Type must be simultaneous fulfillment of the following conditions:

1) the group of AI Individual Type Robots should exists;
2) the presence of an inherited / learned *kinship characters* and *concept* of "one rank" or hierarchical one;
3) the presence of an inherited / learned concept of *group interest.*
4) the presence of an inherited / learned "*process / plan*" to reach the objectives on 3).
5) the presence of an inherited / learned «*incentive / penalty*» (may be several types) to enforce members of the group to follow the *Group Priorities*».

The author is sure that all functions mentioned above can be modelled with known approaches but the real models are out of the represented article.

Надійшла до редакції 29.11.2016

